Augmented Reality as a Potential Tool for Early Detection of Alzheimer's Disease: A Pilot Study

1st Assankhan Amirov
SEEVR Lab, Department of Computing
Goldsmiths, University of London
London, United Kingdom
aamir002@gold.ac.uk

2nd Andrea Castegnaro
Institute of Cognitive Neuroscience
University College London
London, United Kingdom
andrea.castegnaro.15@ucl.ac.uk

3rd Katarzyna Rudzka
Institute of Cognitive Neuroscience
University College London
London, United Kingdom
katarzyna.rudzka.20@ucl.ac.uk

4th Neil Burgess
Institute of Cognitive Neuroscience
University College London
London, United Kingdom
n.burgess@ucl.ac.uk

5th Dennis Chan
Institute of Cognitive Neuroscience
University College London
London, United Kingdom
dennis.chan@ucl.ac.uk

6th Xueni Pan
SEEVR Lab, Department of Computing
Goldsmiths, University of London
London, United Kingdom
x.pan@gold.ac.uk

***Abstract*—Virtual and Augmented Reality (VR and AR) technologies are emerging tools in medical research, and may have particular benefit for early detection of Alzheimer's disease (AD). Consistent with the entorhinal cortex (EC) being the first brain region affected in AD, and the postulated role of the EC in spatial navigation, studies have shown that patients with early AD, prior to dementia onset, are preferentially affected on a VR task of navigation, with the navigational impairment driven by errors in angular estimation. The emergence of consumer grade AR devices provides a novel option for assessing angular estimation which may potentially be more suitable for use in routine diagnostic practice than fully immersive VR. This initial study aimed to design a simplified angular replication methodology and to ascertain the ease of use of AR technology. Twelve volunteers tested the application that required them to replicate an encoding angle using virtual markers within a real world environment. All participants successfully completed the 28 trials consistently within a similar timeframe. Spatial data from the headset provided valuable insights into performance, and questionnaire feedback indicated a positive user experience with minimal simulation sickness. This work demonstrates that angular estimation is testable with AR devices and that AR-based tasks have high user acceptability, highlighting the potential for AR in clinical practice for diagnosis of early AD.**

***Index Terms*—Augmented Reality**

I. Introduction

The advent of new potentially disease-modifying drugs for the treatment of Alzheimer's disease (AD) has increased interest in the detection of AD in its earliest stages, when such treatments may have greatest potential for delaying or even preventing progression to dementia [1]. However, current diagnostic methods are limited either in terms of their suitability for early detection or in their potential scalability given the high prevalence of AD in the aging population. The legacy pen and paper cognitive tests used currently in memory clinics lack both sensitivity and specificity for pre-dementia AD.By comparison, biomarker-based tests examining the presence of the amyloid and tau molecular pathologies associated with AD, such as PET brain scanning or examination of cerebrospinal fluid via lumbar puncture, are not suitable for widespread use given their cost, invasiveness and very limited availability.

A potential solution to this problem may lie in the use of novel tests probing the functions of the brain regions affected in the very earliest stages of AD. Neuropathological studies [2] have shown that, the entorhinal cortex (EC), resident within the brain's medial temporal lobe, is the first cortical region to exhibit neurodegeneration with accumulation of pathological hyperphosphorylated tau protein which spreads trans-synaptically from the EC to the hippocampus and from there to the neocortex.

The advent of commercially available immersive Virtual Reality (VR) headsets provides an opportunity to test EC-based navigation in people with early AD in a way that is more ecologically valid than desktop-based navigation tests that do not involve actual locomotion. Furthermore, the design of a test that is largely independent of language function helps address the confounds that hamper application and interpretation of legacy cognitive tests. When applied to patients with mild cognitive impairment (MCI), in which affected individuals have cognitive impairment but not at a severity to warrant a diagnosis of dementia, VR testing of path integration representing a form of navigation associated with EC grid cell function, was able to differentiate those MCI patients with underlying AD with higher classification accuracy than a basket of current gold standard cognitive tests used in clinical and research practice, including tests of verbal and nonverbal memory [3]. More recently, the same VR test has been used to show that asymptomatic middle aged individuals at high risk of AD (e.g. family history, ApoE4 genotype or physiological risk) are selectively impaired on path integration prior to impairment in other cognitive domains including memory and other spatial behaviors such as egocentric spatial orientation, indicating that altered path integration may represent the transition point from at-risk to disease onset in AD [4].

More recent work has modeled the path integration error in

the MCI patients studied by Howett et al [3], breaking this down into linear and angular path integration and considering in turn encoding (internal representation of linear and angular self-motion), calculation (integration of encoding information to determine inbound path), and production (translating the intended inbound path into physical action). Such modeling revealed that the path integration error in MCI patients with underlying AD was driven by overestimation of angular distance [5], and as such is in keeping with the observation by Newton et al [4] that angular error was the main determinant of the navigational error in people at increased risk of AD.

These clinical findings, and the identification of angular error as the primary contributor to the navigation error in early AD, raises the possibility that this aspect of spatial behavioral impairment may be captured using latest generation AR headsets that may be more suitable than fully immersive VR for usage in routine clinical practice given the need for lower user burden and task complexity in such contexts. This paper focuses on the design and initial application of angular replication tasks using an AR device. We also present the preliminary result from our initial pilot studies with healthy participants. Walkthrough video: https://youtu.be/Y8bcYrvpzvY.

II. Background

A. Place cells, navigational accuracy, and the human hippocampus

The study of spatial memory, particularly within the realm of neuroscience, has long been fascinated by the discovery of place cells in the hippocampus [6]. These specialized neurons play a pivotal role in spatial navigation by firing in response to specific locations within an environment. The EC, an adjoining region, boasts grid cells, which respond to an animal's location based on a hexagonal firing pattern. Researchers have often likened the function of the EC and the hippocampus to the creation of a cognitive map. While grid cells offer a spatial metric, place cells serve to identify distinct landmarks within that space. By synchronously operating, these regions present a malleable cognitive map and memory of a space, capable of adjusting based on changes in the environment [7]. Adaptability was demonstrated in a study where place cells shifted their firing patterns upon the introduction of animals to a new environment. Similarly, changes in the environment's geometry prompted adjustments in grid cell firing. This process was also observed when an animal was navigating a 3 dimensional grid like structure, indicating that the positioning system works in all vectors and is much more sophisticated than previously thought [8].

Central to understanding how AD develops is recognizing its effects on specific brain regions, specifically the previously mentioned EC and the hippocampus. EC is often among the first brain regions to exhibit signs of tau pathology. Tau proteins, under normal circumstances, stabilize microtubules in neuronal cells. However, in AD, they become aggressive and begin to build up, forming tangles. These tangles disrupt nutrient transport within neurons, eventually leading to cell death [9]. Tau proteins play a crucial role in AD, especially within the EC and hippocampus. A study by Calignon et al. used a mouse model to show that tau pathology can move from one group of neurons to another, emphasizing the importance of identifying the initial tau build up in the EC before it begins to spread [10]. Targeting the build up in the early stages can significantly mitigate the effects of AD. These findings underscore the importance of place cells and the EC in spatial memory and navigation, as well as the urgency of identifying any anomalies in their function for early AD diagnosis.

B. Spatial Navigation Tasks in Virtual Environment

VR presents an intriguing avenue for studying spatial navigation and has been utilized to test many different aspects of the human brain, from human interaction to object memory. With many experiments proving that the participants were immersed in the task and produced tangible results compared to real world tasks. Indicating that the VR technology is a viable tool in the research of human cognition, among other possibilities [11]. An experiment involving 49 participants (19 with mild cognitive impairment) utilized an immersive VR path integration task to differentiate the participants with mild cognitive impairment [3]. This task was administered using the HTC Vive iVR kit, allowing participants to physically navigate a 3.5 x 3.5 m space. Within the Path Integration task the participants were asked to navigate a triangular path, guided by one visible marker at a time, and then asked to return to the first marker without any ques. The task required that the participants keep track of their movement and rotation, as the VR environment lacked any close landmarks the participants could orient from. The only thing visible to the participants was the far off landscapes, such as mountains and ranges. The task was designed to measure the EC based navigation and the error rate between the right position and the position that the participant thought was the start. Prominently, the MCI group underperformed when juxtaposed with the control group. Moreover, the study's VR task, rooted in the EC's navigation function, proved more useful at identifying early cognitive impairments than conventional neuropsychological tests. This study champions the argument for employing VR-based navigation tasks, spotlighting their potential in the early detection and intervention in Alzheimer's disease [12].

C. AR for Spatial Navigation Tasks

AR shares a lot of the benefits of immersive VR whilst providing further simplifications which may render it a tool particularly suitable for clinical settings. As with VR, the real-world movements permissible within, AR tasks allows gathering of physiological self-motion information including proprioception (feedback from mechanosensory neurons), vestibular feedback (information related to a sense of balance), and motor efference copy (internal copy of the motor commands). The possible advantage of AR is that the passthrough mode may increase the usage comfort and safety by remaining embedded in the real environment in populations who have not been exposed to these technologies, in particular when there is a limited time to administer the tests in clinical settings. Spatial

memory tasks could be designed using AR and allow for a more contextualized and personalized experience for the participants and their environment [13] whilst in a relatively lightweight headset which is easy to setup owing to the inside-out tracking. Existing literature have proven AR being valuable in application in researching and training memory of the participants [14]. A study comparing spatial memory tasks in VR and AR revealed participants' inclination towards AR, citing its ease, immersion, and enjoyment [15]. This shift towards AR isn't merely preferential; participants also exhibited improved memory performance. The study indicates that AR, by presenting a more realistic, dynamic environment, stands as a promising tool, not only for spatial memory research but also for therapeutic interventions targeting spatial memory impairments.

III. AR Task Design

A. Angular Replication Task

In an attempt to provide a simple metric of angular path integration, this task measures how well participants remember and subsequently replicate simple turns. In each task, participants complete a guided turn between two reference points. Subsequently, they are asked to continue turning in the same direction until the believe they have turned by the same amount, and confirm their response with a trigger button. As shown in Figure 1, a trial begins with the participants facing a reference mark M1, and rotating until M2, resulting in a physical rotation of X1, which they should would aim to replicate by rotating to MA where X1 = X2. Replication error on this task is reflected as Delta Error (DE) and computed as the angle between their physical location (M3) and where they should have rotated to (MA). In each trial, the originally encoded angle and its replication are in the same direction (i.e., either both clockwise or anticlockwise). This is to prevent the use of simple landmarks as a response strategy.

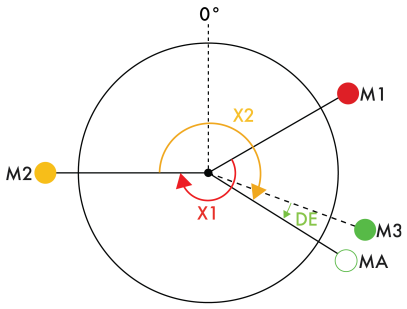

Fig. 1. Angular task.

B. Angles

The angles that the patient has to replicate are selected based on the following principles:

- Not dividable by 90 or 45, as those rotations are instinctive and commonly performed.

Modules	Angles						
A1	150	330	210	120	240	60	300
A2	330	210	60	120	150	300	240
B1	240	60	120	330	210	150	300
B2	60	150	330	240	120	300	210
C1	210	300	240	60	120	150	330
C2	120	240	330	60	300	210	150

Fig. 2. Tested angles. Highlighted in red are anticlockwise trials

- Bigger then 60 degrees, as the replication by people tends to lose accuracy at such short rotations. Similarly, the angles should not be larger than the full 360 rotation to prevent people from completing a full circle as it negates the need to remember the rotation.
- Larger angles are preferred, as the patient would have to utilize more of their spatial memory.

Based on those rules, 7 angles were chosen: 60, 120, 150, 210, 240, 300 and 330 degrees [16]. These angles could also be utilized in a clockwise and anticlockwise rotation, generating 14 testable angles. A total of 3 modules (A,B, & C) including 42 angles were designed, as shown in Figure 2. Those highlighted in red indicate anticlockwise direction. Each module contains an even number of anticlockwise and clockwise angles. For our user study, each user went to the first two modules (A & B), resulting in 28 trials.

C. Augmented Reality Method

Translating the angular task to the AR application has produced some challenges, but also design possibilities. As shown in Figure 3: Left, the participant is asked to stand inside specifically marked zone shown as a circle under the patient, and then begin by facing the first starting marker (1), which is indicated in red. This red marker will be floating in space in front of the participant at their eye level. Once standing in the middle of the marked zone and looking directly at the red mark, the task will be activated and the red mark will inflate in scale and play a activation sound. The participant doesn't know the exact position of any of the markers in advance, as they are hidden until directly looked at (so that the participant does not use peripheral vision to predict where they will end up). Once the starting marker has been activated it will prompt the patient to begin rotating in the direction indicated by UI arrows, either clockwise or anticlockwise. The application instructs the patient to turn their whole body, and not just the head. Continuing the rotation, the patient will eventually encounter the yellow encoding marker (2), which indicates to them the end of the rotation that they must replicate. Similarly to the starting marker (1), the participant will not see the encoding marker (2) until they look directly at it, this also prompts the participant to rotate slowly in order to catch each marker. They will hear an activation sound once it's activated. The participant is then prompted to continue the rotation in the same direction and mark the position at which they think that the angle has been successfully replicated. Once the participant is happy with their selection, and confirms it using the aim ball,

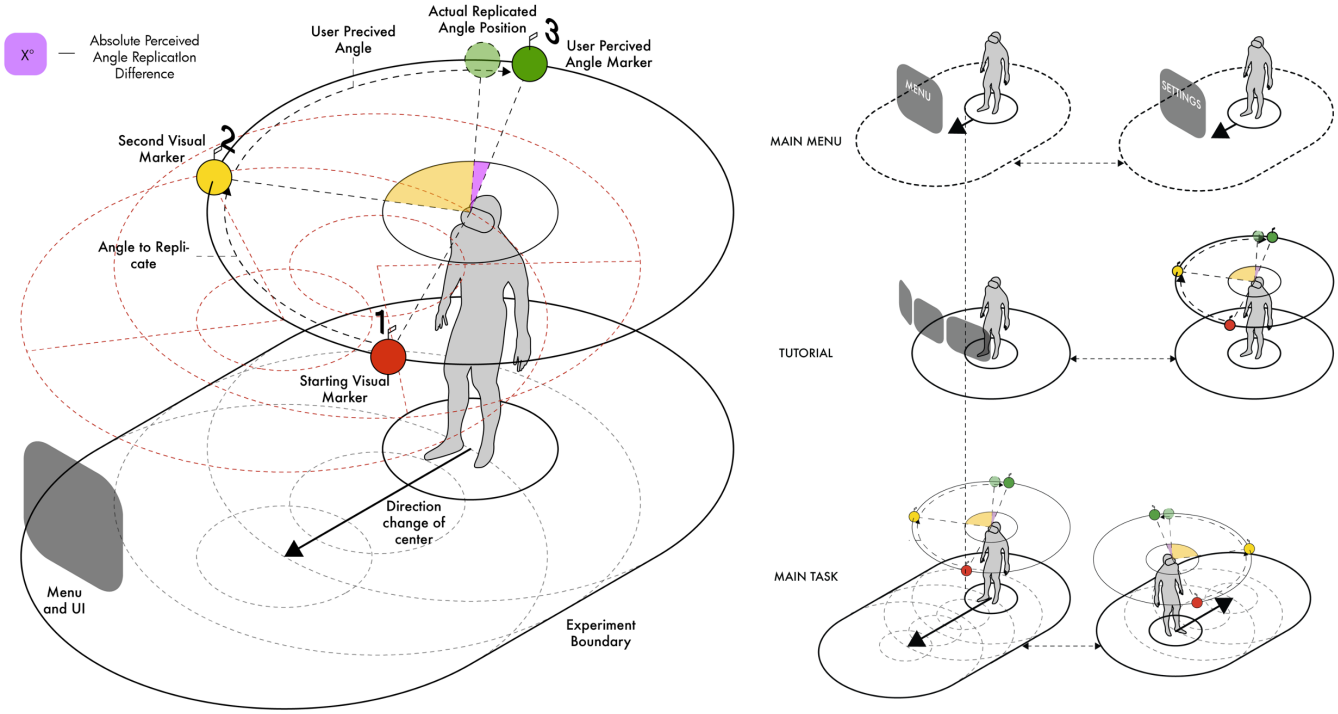

Fig. 3. Left: Diagram of the AR Angular Task Design; Right: UI and Navigation

they press the left-hand trigger on the controllers, at which point a green participant angle mark (3) will appear, and with the familiar sound, it finalizes a single trial.

In an ideal scenario, the participant should spend less than 30 or 40 seconds per angle, allowing them go through them in rapid succession. After the each trial, the UI will pop up a message reminding the participant to locate the new center of the task and promptly move towards it. This is done to allow the participants to mentally differentiate different angle tasks, and have a mental break between the rotations. The process continues until the participant has completed all of the trials.

D. Hardware

We use the Passthrough mode of the VIVE Elite XR device. It is a standalone headset, allowing participants to move around the space freely. Its passthrough video feed of the headset provides a clear and immersive image of the real world, while its depth sensors ensure an accurate sense of depth. This combination of virtual and real-world imageries can be particularly beneficial for medical research.

E. Gaze Interactions

The manner in which the participant orients and faces the objectives is one of the most important elements of the task. We use the Unity raycast function to obtain participant's head position and rotation. Eye tracking was not yet available for Elite XR at the time of our experiment (and has only come to the market at the time of writing).

F. Data Logging

A CSV file is created automatically for each participant, logging the ID of events, the time stamp, participants' head position and rotation, and the trial info (angle, anticlockwise or not, delta error). A total of four different events were considered: (1) Reposition to the new floor circle; (2-4) Activation of the first/second/third marker.

G. Spatial Audio

Since the application is utilized in real-world space and the participant moves around, spatial audio was used to enhance their immersion and make the markers feel more present in the scene. This way, when a participant looks or walks away from the target that they should be focusing on, the sound of the narration gets quieter, prompting the participant to move back closer to hear it better.

H. Direction of Turning

The application tries to guide the participant to rotate towards the required direction, whether clockwise or anti-clockwise, and measures the delta change of the rotation. Two arrows are displayed, one indicating the correct direction, the other one the incorrect one (Figure 4, bottom right). The level of opacity of each arrow changes as participants rotates - the one representing their current rotation becomes more solid over time, and vice versa. We found this an intuitive way to guide participants towards the correct direction.

The AR layout of the menu are also spatial, with each layer building on top of the other. This can be seen in Figure 3: right, which illustrates the different scenes of the application.

The first level is the Menu Scene, it outlines the zone of the testing and the starting position. The Menu UI screen has three components, Tutorial, Start, and Settings. The first two buttons take the user to the corresponding levels, and the last one allows the experimenter to adjust settings and input participant ID which generates the CSV file for data logging.

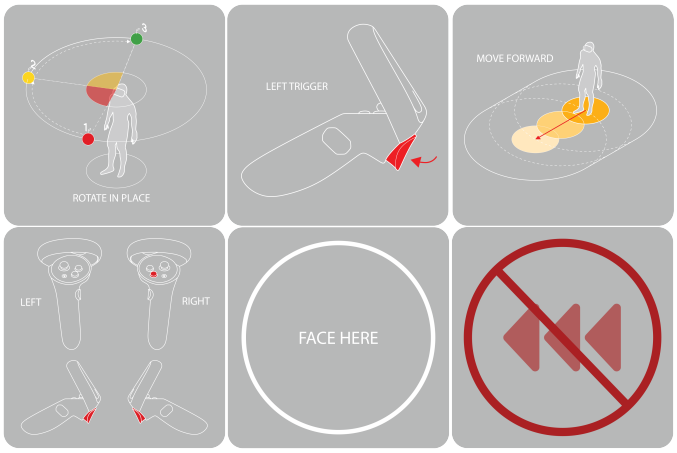

Fig. 4. UI Icons.

I. Tutorial

Whilst the participants would be accompanied by a clinician most of the time, the tutorial provides a comprehensive guide for the participants to do the task in a safe environment, allowing them to make mistakes and adjust to the movement in Augmented Reality. The participant was guided at each step by an audio-generated voice, indicating in detail what and when the participant should do. The audio was generated using Azure Voice Generation AI service, it allowed for control of intonation, and pauses and produced easily transferable audio clips for the Unity Application.

Once these have been completed, the participant is directed to return to the main menu and begin the actual trials. A walk-through video can be found here: https://youtu.be/Y8bcYrvpzvY.

IV. User Evaluation

A. Participants

The participants were a mix of general public volunteers and associated researchers. There have been 12 participants, the majority of them are 26-30 (9/12), two between 18-25 and one 31-35. Nine were male and three female. While a wider study focusing on the populations with MCI markers was also planned, this particular user study did not attempt to recruit participants on any specific criteria in terms of mental cognition state. The participants were included in the study based on their willingness to participate, they were not paid for their time. Many of them were reached through social media, wider University emailing lists, and word of mouth. The entire process of questionnaires and the task took an average 31m and 05s. Some trials were conducted at various workshops and conferences, however, the data from those trials was not included in the data analysis here as it was not done under experimental settings. The feedback from those additional trails was included in the later discussion section.

B. Procedure

On the participants' arrival to the lab space, they were greeted by a research assistant and asked to fill in a questionnaire on a laptop. After they were asked to move to the center of the open space and equip the HTC VIVE XR Elite (Figure 5 B&C). The researcher assisted the participants with the headset, such as adjusting the lens setting for the participants with glasses prescription and familiarizing them with the different elements of the headset. The participant was given time to adjust to the view of the augmented reality, as they could see themselves and the space around them, but now with technical overlay. Once the participant was adjusted to the movement in the headset, they were allowed to begin the tutorial and the tasks. The researcher assisted the patient by monitoring their progress through a tablet streaming device which was paired to the headset, however, the researcher could not adjust any of the settings of the experiment from the streaming device.

After the completion of all of the required tasks, the researcher helped the participant to remove the headset, and a small break was allowed before the participant was asked to complete the post-trial questionnaire. At the end of the second questionnaire, the experiment was concluded. The data from the questionnaires was automatically tiled up by the Microsoft Forms, and the data from the angular task trials was downloaded from the headset at the end of the day. The headset automatically generated a CVS file for each participant to store their data.

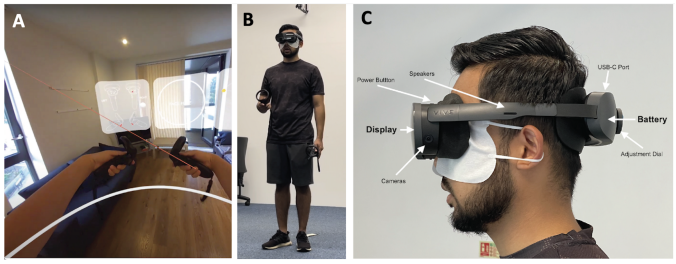

Fig. 5. A: AR View from participants' perspective; B & C: AR HMD

C. Space Setup

The user study was conducted in our VR research lab in order to control for the environment and the participant experience. Our VR Lab is a 6m x 8m open space, with tables and computers arranged around the perimeter of the room.

D. Questionnaires

The questionnaire was designed in order to gauge the participant's familiarity with the AR and VR technologies, their immediate emotional and physical state, their memory abilities, and their participant experience. First of all, we

used the standard Simulated Sickness Questionnaire (SSQ) [17] both before and after the experiment to understand how people felt in the Augmented reality environment. As our experiment includes a lot of turning around it is possible that the participants developed simulation fatigue and the questions could help us understand the extent to which it does so.

Secondly, we used part of the Multifactorial Memory Questionnaire (MMQ) [18] to determine the perceived memory abilities, with the answers being scored to determine their baseline. Although their memory was not the primary focus of this study, these questions made the participants reflect on their spatial memory and could provide insights into the better design of the experiment for people with worse perceived memory.

Finally, we used an adapted version of the Virtual Reality Neuroscience Questionnaire (VRNQ) [19] to evaluate users overall experience. We included the first three categories (user experience, game mechanics, and in-game assistance) to better understand participants' experience and they could be improved. The data from the questionnaire was anonymized as the name of the participants was never asked, only the Participant ID. It was conducted through the Microsoft Forms platform on the University account.

V. Results

The main data gathered per each participant is the CSV file containing all of their movement and interaction data from the task (see Section III-F) and the questionnaire response that they have filled before and after the task. The focus of this paper is the usability of the application and the participants' experience rather than the angular task error rate itself. We first present the questionnaire results, followed by behaivoural result, where we hope to use for future reference in our overall research project.

A. Questionnaire Results

1) VR/AR Usage Questionnaire: Before the experiment each participants filled a questionnaire about their general usage and familiarity of V/AR, and its usage in the clinical setting. Of the 12 participants, only 2 have not used a VR or AR headset before. This is to be expected as the sample size of graduate students and young professionals in London would have encountered the technology as it has become more mainstream in everyday settings. When asked about their level of familiarity of VR and AR technology from a Likert scale of 0-10, nine chose above the median (> 5), with an average answer of 6.8 ± 2.9. When asked about their familiarity with the use of V/AR technology in the medical setting, only four chose above the median value, with an average answer of 4.8 ± 3.3. Half of the participant reported that they own a V/AR device, only two of them said they have used V/AR in a clinical setting.

2) VRNQ: For the 12-item adapted VRNQ, we used a 0-10 Liker-scale. This is because we would like to later analyse the result using the Net Promoter Score [20] and [21], a popular concept used in market research. We calculated the average score for each participant for User Experience (UX) over 5 items, Game Mechanics (GM) 2 items, and In-Game Assistance (GA) 5 items. Overall, our applications received high ratings for all these three categories (UX: 8.0 ± 1.1; GM: 7.8 ± 1.4; GA: 8.1 ± 1.3). Our correlation analysis showed that, there were a significant positive correlations between UX and GA ($r^2 = .52$, **p = .008**), and between GM and GA ($r^2 = .65$, **p = .002**), but not between UX and GM ($r^2 = .30, p = .065$).

3) Pre and Post SSQ: Participants were given the SSQ questionnanire twice, before and after their AR experience. As expected, there is a significant correlation between pre and post SSQ ($r^2 = .498$, **p = .01**). Shapiro-Wilk test found our preSSQ not normally distributed (p = .010). We used the non-parametric Wilcoxon signed-rank test to compare pre and post SSQ. Although post SSQ is higher than pre SSQ (preSSQ: $3.17 \pm 3.95.14$, postSSQ: 5.00 ± 3.98), the difference is not significant ($z = -1.853, p = .064$).

As simulation sickness is also part of user experience, we run a correlation analysis between diffSSQ (postSSQ - preSSQ) and the VRNQ elements. We found a significant *negative* correlation between diffSSQ and GA ($r^{=}.49$, **p=.011**). This suggests the more the application made the participant felt dizzy, the worse they rated the Game Mechanics.

There is also a negative correlation between UX and diff-SSQ, and GM and diffSSQ, but the correlations were not significant (UX vs diffSSQ: $r^{=}.12, p = .26$; GM vs diffSSQ: $r^{=}.28., p = .08$).

B. Behavioural Data

The behavioural readout of this angular AR task is the Delta Error (DE), an absolute difference between the participant's actual physical response and the rotation required to ideally replicate the angle (Section III-A. DE were computed for all participants for each of the 28 completed angles. To provide an error metric scaled to the magnitude of each angle, we normalised the errors by dividing DE for each trial over the correct answer (e.g., if participants turned 150 when they were supposed to turn 120, the noramlized DE would be $(150-120)/120 = 0.25$). This metric was used as participants' performance indicator to address the following hypotheses:

- **H0**: There is no difference between clockwise trials and anticlockwise trials.
- **H1**: Participants' performances would improve over time.
- **H2**: There is a positive correlation between participants' perceived memory abilities and their actual performance in our tasks.
- **H3**: Participants who made more mistakes found it more difficult, and would have also spent longer on the tasks.

H0 here is more of a validation hypothesis, to validate our methods. And as the number of clockwise and anticlockwise trials are equal, we are then able to test H1-3. In the following, we first present the overall time participants spent on the task, and then our data anlysis on our H0-3.

1) Overall Time: We logged participants time when they started each trial, and calculated the estimation of the overall time spent in the AR environment (for the first 27 trials, as we do not have the time when they completed the 28th trial). On average, participants spent 12.5 minutes on those trials, with the fastest one spending 9.5 minutes, and the slowest 20.5 minutes.

*2) Clockwise and Anticlockwise (**H0**):* Half of our trials are clockwise and half anticlockwise trials, and we do not expect any differences between the two. For each participant, we calculated the average of their normalized DE for all their clockwise trials (14 trials) and anticlockwise trials.

Our data passed the test of normality (Shaprio-Wilk test, clockwise: $p = .41$; anticlockwise: $p = .37$). We used the repeated ANOVA to compare the average of the normalised error for all the clockwise trials and anticlockwise trials, for each participant. Participants normalised delta error was similar in the clockwise trials ($.12 \pm 0.04$) than in anticlockwise ones ($.15 \pm 0.06$), suggesting that participants performed at the same level between clockwise trials and the anticlockwise ones ($F(1, 11) = 2.49, p = .143, \eta^2 = .185$). As there was no difference in DE between the clockwise and anticlockwise turns for the same angles, we collapsed and analysed our results from both types of trials together for the following analysis.

*3) Performance Over Time (**H1**):* In order to understand participants' performance over time, we split the 28 trials into the first half and the last half, with 14 trial in each. As all 7 angles appear twice, once clockwise and once anticlockwise, the level of difficulties is matched across both halves.

For each participant, we calculated the average of their normalized DE for their first (14 trials) and second half. As our data was not normally distributed (Shaprio-Wilk test, first_14 $p = .032$, second_14 $p = .026$), we used the Wilcoxon signed-rank test to compare the averages. Participants normalised delta error did not significantly differ between the first (0.142 ± 0.040) and the second half (0.116 ± 0.538) , ($Z = -1.805, p = .071$).

We then compared the first 7 trials with the last 7 trials, as again, they have the same set of numbers so should be similar in their level of difficulties. Test of normality was rejected (Shaprio-Wilk test, first_7 $p = .029$, last_7 $p = .002$) so we again we used the Wilcoxon signed-rank test. Participants' performance was worse at the beginning (0.154 ± 0.040), but improved towards the end (0.106 ± 0.064). The difference between these two was significant ($Z = -2.276$, **p = 0.023**).

*4) Performance and Perceived Memory Ability (**H2**):* We were interested in whether participants performance is linked to their perceived memory ability, as measured by MMQ. We run a correlation analysis on the overall performance (average of all 28 trials of normalized DE) and MMQ, and found no correlations between the two ($r^2 = 0.002, p = .900$). We also run a correlation analysis on the overall time participants spent on the task and MMQ, but again found no correlations between the two ($r^2 = 0.052, p = .477$).

*5) Delta Error Rate and Time Spent (**H3**):* Lasty, we addressed the hypothesis that individuals who had worse performance (as indexed by normalised DE) were the ones who struggled with the task, and thus would have spent more time on it. However, contrary to our hypothesis, there is a significant *negative* correlation between time and normalised DE ($r^2 = 0.052$, **p = .041**). This suggested that those who spent longer on the task in fact performed better.

VI. Discussion

First of all, our VRNQ result showed that our AR application was very positively received, with all three aspects of the VRNQ. This means that our implementation of the AR tasks were successful in terms of user experience. When it comes to SSQ, although not significant, our data still indicates some degree of increase in simulation sickness, with 8 participants showing an increase in the values. The XR Elite is still projecting only a stitched image onto eyes, with a limited resolution and visual distortion of close up objects, the technology still has a lot of room for improvement. However, it is also useful to note that the result from the SSQ Questionnaire ranges from 0 to 64 (16-item Likert Scale from 0 to 4), thus our post SSQ of 5.00 could be considered negligible.

Interestingly, we found negative relationships between the user experience and participants change in SSQ, indicating the lower the change in SSQ, the higher the participants rated the user experience. This means that if we could further examine aspects in our application to reduce simulation sickness, our user experience could be further improved. The causal effect of this relationship needs to be examined in further studies.

Our behavior data first confirmed that all participants were able to complete the task within very reasonable time, and that there were no differences between participants' performance in clockwise and anticlockwise trials (**H0**). We also saw a learning effect, with participants performing worse in the beginning as comparing to the end, supporting **H1**. Further, we were able to provide evidence that this effect diminished after the first 7 trials. More studies need to be conducted to determine what would be the optimal trial numbers for the learning effort to wear off.

We found no relationship between their perceived memory ability and actual performance, as measured by either time spent or average of their normalised DE. So our **H2** is not supported. However, we did find a negative relationship between the two performance indicators themselves: those who spent longer completing the task produced less errors, rejecting **H3**. We think this could be caused by level of conscientiousness of the participants themselves, rather than the task. In future, we could include the big-five personality questionnaires to control for this factor.

We conducted a small qualitative interview after the experiment and received very positive feedback of the overall experience, and in particular in regards to the UI and Tutorial aspects of the application. There were a couple instances of the headset malfunction due to small battery capacity and overheating during extended use, as well as border zone issues

when the headset lost tracking, but all of the participants were still able to complete all of the tasks. Overall, we found several benefits of using the VIVE Elite XR: it has been easy to setup, provides high accuracy in data recording, and has adjustment to suit people with different eyesight. These elements provided a comfortable experience to the participants, which is incredibly important in medical screening, as uncomfortable test cases could produce unreliable results when the patients would want to finish faster or not be motivated to do the task right. Also, the portable and compactness of the headset allows it to be utilized in any environment.

We also asked the participants after the experiment about their method for gauging their angle of rotation. Many relied on their environment and certain interior landmarks to navigate. As mentioned before, the right angle nature of the buildings and our environment provides contextual clues to the participant where the right angles are, and that could be mitigated by performing the task in an open space or at a random angle to the room orientation so that the room grid becomes irrelevant. Some other answers included internal counting; the participants would count out the numbers in a set interval in order to measure the angels. In future, it might be useful to instruct the participants to count backwards from 10 to 1 on repeat whilst doing the task, in order to mitigate internal vocalization assistance.

VII. Conclusions

The findings of this study provided useful insights into the potential usability of A/MR technologies in medical screening, as the findings demonstrated positive user experiences and limited simulation sickness. The technology and design principles for the tasks could also be utilized in the broader field of cognitive neuroscience. This research will continue as part of the overall research project, with lessons learned from this study paving the way for an improved version of AR Angular Replication Task. Overall, our study demonstrated the potential of AR in spacial cognition testing and as a clinical screening tool for early AD diagnosis beyond the capacity of traditional 2D or Computer memory tests.

Acknowledgment

Pan is funded by the AHRC Project "Immersive, Innovative, and Interactive Experience" AH/T011416/1.

References

[1] J. R. Sims, J. A. Zimmer, C. D. Evans, M. Lu, P. Ardayfio, J. Sparks, A. M. Wessels, S. Shcherbinin, H. Wang, E. S. M. Nery *et al.*, "Donanemab in early symptomatic alzheimer disease: the trailblazer-alz 2 randomized clinical trial," *Jama*, vol. 330, no. 6, pp. 512–527, 2023.

[2] H. Braak and K. Del Tredici, "The preclinical phase of the pathological process underlying sporadic alzheimer's disease," *Brain*, vol. 138, no. 10, pp. 2814–2833, 2015.

[3] D. Howett, A. Castegnaro, K. Krzywicka, J. Hagman, D. Marchment, R. Henson, M. Rio, J. A. King, N. Burgess, and D. Chan, "Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation," *Brain*, vol. 142, no. 6, pp. 1751–1766, 2019.

[4] C. Newton, M. Pope, C. Rua, R. Henson, Z. Ji, N. Burgess, C. T. Rodgers, M. Stangl, M.-E. Dounavi, A. Castegnaro *et al.*, "Entorhinal-based path integration selectively predicts midlife risk of alzheimer's disease," *Alzheimer's & Dementia*, 2024.

[5] A. Castegnaro, Z. Ji, K. Rudzka, D. Chan, and N. Burgess, "Overestimation in angular path integration precedes alzheimer's dementia," *Current Biology*, vol. 33, no. 21, pp. 4650–4661, 2023.

[6] J. O'Keefe, N. Burgess, J. G. Donnett, K. J. Jeffery, and E. A. Maguire, "Place cells, navigational accuracy, and the human hippocampus," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 353, no. 1373, pp. 1333–1340, 1998.

[7] N. T. Robinson, L. A. Descamps, L. E. Russell, M. O. Buchholz, B. A. Bicknell, G. K. Antonov, J. Y. Lau, R. Nutbrown, C. Schmidt-Hieber, and M. Häusser, "Targeted activation of hippocampal place cells drives memory-guided spatial behavior," *Cell*, vol. 183, no. 6, pp. 1586–1599, 2020.

[8] R. M. Grieves, S. Jedidi-Ayoub, K. Mishchanchuk, A. Liu, S. Renaudineau, É. Duvelle, and K. J. Jeffery, "Irregular distribution of grid cell firing fields in rats exploring a 3d volumetric space," *Nature neuroscience*, vol. 24, no. 11, pp. 1567–1573, 2021.

[9] T. Gómez-Isla, R. Hollister, H. West, S. Mui, J. H. Growdon, R. C. Petersen, J. E. Parisi, and B. T. Hyman, "Neuronal loss correlates with but exceeds neurofibrillary tangles in alzheimer's disease," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 41, no. 1, pp. 17–24, 1997.

[10] A. De Calignon, M. Polydoro, M. Suárez-Calvet, C. William, D. H. Adamowicz, K. J. Kopeikina, R. Pitstick, N. Sahara, K. H. Ashe, G. A. Carlson *et al.*, "Propagation of tau pathology in a model of early alzheimer's disease," *Neuron*, vol. 73, no. 4, pp. 685–697, 2012.

[11] A. Castegnaro, D. Howett, A. Li, E. Harding, D. Chan, N. Burgess, and J. King, "Assessing mild cognitive impairment using object-location memory in immersive virtual environments," *Hippocampus*, vol. 32, no. 9, pp. 660–678, 2022.

[12] D. E. Adamo, E. M. Briceño, J. A. Sindone, N. B. Alexander, and S. D. Moffat, "Age differences in virtual environment and real world path integration," *Frontiers in aging neuroscience*, vol. 4, p. 26, 2012.

[13] A. Niknam, "An augmented reality mobile game design to enhance spatial memory in elderly with dementia," in *6th International Conference on Computer Games; Challenges and Opportunities (CGCO2021)*, 2021.

[14] M.-C. Juan, M. Mendez-Lopez, E. Perez-Hernandez, and S. Albiol-Perez, "Augmented reality for the assessment of children's spatial memory in real settings," *PloS one*, vol. 9, no. 12, p. e113751, 2014.

[15] S. Maidenbaum, A. Patel, I. Garlin, and J. Jacobs, "Studying spatial memory in augmented and virtual reality," *bioRxiv*, p. 777946, 2019.

[16] E. R. Chrastil and W. H. Warren, "Rotational error in path integration: encoding and execution errors in angle reproduction," *Experimental brain research*, vol. 235, pp. 1885–1897, 2017.

[17] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[18] K. T. Shaikh, E. L. Tatham, J. B. Rich, and A. K. Troyer, "Examining the factor structure of the multifactorial memory questionnaire," *Memory*, vol. 29, no. 2, pp. 255–260, 2021.

[19] P. Kourtesis, S. Collina, L. A. Doumas, and S. E. MacPherson, "Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology," *Frontiers in human neuroscience*, vol. 13, p. 417, 2019.

[20] F. F. Reichheld, "The one number you need to grow," *Harvard business review*, vol. 81, no. 12, pp. 46–55, 2003.

[21] K. Liyakasa, "Customer experience is critical in net promoter benchmarks," *CRM Magazine*, vol. 16, no. 6, pp. 16–16, 2012.